



## King's Research Portal

DOI:

[10.1093/fampra/cmn025](https://doi.org/10.1093/fampra/cmn025)

*Document Version*

Early version, also known as pre-print

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Slade, M., Gask, L., Leese, M., McCrone, P., Montana, C., Powell, R., Stewart, M., & Chew-Graham, C. (2008). Failure to improve appropriateness of referrals to adult community mental health services: lessons from a multi-site cluster randomized controlled trial. *Family Practice*, 25(3), 181 - 190. <https://doi.org/10.1093/fampra/cmn025>

### Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### Take down policy

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

**Failure to improve appropriateness of referrals to adult community mental health services – lessons from a multi-site cluster-randomised controlled trial**

**Running head: Referrals to adult mental health services**

Dr Mike Slade <sup>1</sup>

Health Service and Population Research Department (HSPRD), Institute of Psychiatry,  
King's College London

Prof Linda Gask

School of Community-Based Medicine, University of Manchester

Dr Morven Leese

Health Service and Population Research Department (HSPRD), Institute of Psychiatry,  
King's College London

Dr Paul McCrone

Health Service and Population Research Department (HSPRD), Institute of Psychiatry,  
King's College London

Carolyn Montana

School of Community-Based Medicine, University of Manchester

Dr Robin Powell

Park Royal Centre for Mental Health, London

Mairi Stewart

Health Service and Population Research Department (HSPRD), Institute of Psychiatry,  
King's College London

Dr Carolyn Chew-Graham

School of Community-Based Medicine, University of Manchester

---

<sup>1</sup> Corresponding author: Health Service and Population Research Department (Box P029), Institute of Psychiatry, King's College London, Denmark Hill, London SE5 8AF

# **Failure to improve appropriateness of referrals to adult community mental health services – lessons from a multi-site cluster-randomised controlled trial**

## **ABSTRACT**

### **Background**

Non-clinical factors impact on decisions about whether to refer a patient from primary care to specialist mental health services. The aim of this study was to investigate whether introducing a standardised assessment of severity improves agreement on referrals.

### **Methods**

Multi-site mixed-method cluster-randomised controlled trial, investigating General Practitioner (GP) referrals from 73 practices (408,839 patients) to 11 Community Mental Health Teams (CMHTs). Intervention group GPs were asked to complete a Threshold Assessment Grid (TAG) rating of mental health problem severity. CMHTs rated referral appropriateness. ISRCTN86197914.

### **Results**

281 GPs made 1,061 mental health referrals. The intervention was only partly implemented, with 25% of intervention group GPs completing TAGs. No difference was found in appropriateness (OR 1.18, 95%CI 0.91-1.53) or secondary outcomes. Post-referral primary care contact rates were higher for the intervention group (IRR 1.36, 95%CI 1.07-1.73). Qualitative data identified professional and organisational barriers to implementation.

### **Conclusions**

Asking GPs to complete a TAG when referring to CMHTs did not improve primary-secondary care agreement on referrals. Low implementation means that uncertainty remains about whether introducing a severity-focussed measure into the referral process is beneficial.

Introducing local protocols to manage demand at this interface may not be successful and more attention needs to be paid to human and organisational factors in managing interfaces between services.

**Key-words:** Primary Care, Community Mental Health Services, Referral, primary-secondary interface

**Failure to improve appropriateness of referrals to adult community mental health services – lessons from a multi-site cluster-randomised controlled trial**

## **INTRODUCTION**

Patients with mental health problems are mostly managed in primary care with patients with more severe mental health problems receiving additional care from specialist secondary mental health services. The four-fold increase between 1971 and 1997 in levels of referral<sup>1</sup> indicates that effectively managing the primary-specialist care interface is a policy priority. An emergent demand management strategy is to triage by severity. Patients with mild to moderate presentations (especially of common mental disorders) are managed in primary care. Patients with more severe mental health problems are managed by specialist mental health services, with input from primary care services varying between countries, but generally including physical health care and some involvement in co-ordinating care.

Ensuring that the right patients are seen by the right part of the mental health system remains problematic<sup>2</sup>. Fewer than 50% of people with common mental disorders will be identified by General Practitioners (GPs)<sup>3</sup> and once identified, a decision has to be made whether to treat solely at primary care level or refer to specialist mental health services, and there is evidence that severity of illness is not the only factor considered. Thus patient gender influences the decision to refer<sup>4</sup>, as does the GP relationship with the patient<sup>5</sup> and the personal threshold for the individual GP<sup>6</sup>. In general, community-based services struggle to retain a focus on the SMI<sup>7</sup> and the imposition of referral guidelines on primary care does little to help<sup>8</sup>.

When people are referred due to factors other than clinical need, efficiency and consequently access is reduced because some people are assessed by the CMHT who do not need to be assessed, and other people who would benefit from specialist mental health care may not be referred. There is disagreement about appropriateness for 20% of primary care referrals to Community Mental Health Teams (CMHTs)<sup>9</sup>, indicating that this is a priority area for innovations to improve access.

We hypothesised that this disagreement is partly due to a differing emphasis on severity – primary care clinicians focussing on supporting their patient in accessing specialist input and de-emphasising comparative severity (because every patient is important), and secondary care clinicians using severity in decision-making both in order to implement policy imperatives and as a non-transparent demand management strategy (*i.e.* raising the severity threshold when capacity is over-stretched). The Threshold Assessment Grid (TAG)<sup>10</sup> is a brief referrer-rated assessment of severity of mental health problems, which has been developed and evaluated for use by referrers to mental health teams.

The aim of this study was to investigate whether introducing a standardised measure of mental health problem severity (TAG) into the referral process improved agreement between primary care referrer and referred-to Community Mental Health Teams on the suitability of the referral for specialist mental health services. A GP practice-level cluster design was used to address within-practice contamination.

## **METHODS**

### *Design*

The study was a mixed-method multi-site cluster randomised controlled trial (RCT). GP practices were randomised, and the unit of analysis was an individual referral of a patient to adult mental health services. The intervention was intended to change GP behaviour, so the design was clustered to reduce the likelihood of contamination<sup>11</sup>. The cluster was chosen as the practice rather than the GP because of the logistical problems which would have arisen if a practice contained both intervention and control group GPs.

Four hypotheses were tested, all at the individual referral level. The primary hypothesis was that the intervention will improve the agreement between the GP and the CMHT on the ‘appropriateness’ of the referral. The secondary hypotheses were that the intervention would make it easier for the mental health team to identify (i) the urgency of the referral and (ii) the most appropriate professional to make the initial assessment, and (iii) be associated with reduced referral discussion time in the CMHT referral meetings. Additionally, the interface and referral process were explored using qualitative methods, and the resource consequences of using TAG were investigated.

Preliminary evidence about effectiveness was available from previous work<sup>9</sup>, which in turn was based on previous theory<sup>12</sup> and modelling<sup>10,13</sup>. The study was intended to be a definitive RCT within the MRC complex intervention evaluation framework<sup>11</sup>. Appropriate ethical and research governance approvals were obtained for each site.

### *Sample and setting*

The sample size for the primary hypothesis was based on 10 mental health teams each receiving 24 referrals per month for 6 months (1,440 in total), of which 65% (936) come from primary care. With a participation rate of 90% this would give 842 referrals from participating GPs, *i.e.* 421 referrals per arm. Assuming an average of 5 referrals per GP, this sample size is equivalent to an approximate effective sample size of 234 per group, assuming an intraclass correlation for referrals from individual GPs of up to 0.2 leading to a design effect of 1.8. (The *a priori* power calculation was done at the design stage when the randomisation was envisaged as being by GP. When it was decided that the randomisation was to be by practice, it was decided to retain a sample large enough to allow for GP clustering also, as this was likely to have a larger effect than practice clustering. A deliberately conservative estimate for the GP intraclass cluster correlation was chosen. In fact the GP cluster design effect was overestimated but the design effect assumed (1.8) turned out to be approximately correct for the practice intraclass cluster correlation actually obtained (1.84 with an icc of 0.06 and 15 referrals per practice). Previous work found 20% of referrals were rated as inappropriate<sup>9</sup>, so this sample size would allow a difference to be detected in the proportion of inappropriate referrals between the groups of 10% (*i.e.* a 50% drop from 20% to 10% in the intervention group) with a power of 83%, using a significance level of  $p=0.05$ . In the event, the trial was extended to 9 months and 11 teams to obtain the needed sample size.

The sites were chosen to meet three goals: (i) to have a nationally representative population; (ii) to include high deprivation inner-city areas in which primary care services are less comprehensive, leading to increased pressure on the primary-secondary interface (since these are the areas where the TAG will have the highest potential impact); and (iii) to allow health economic analysis of the system-wide impact of the intervention on a sizeable area. To meet



these goals, the sites chosen were one complete London Borough (Croydon) comprising eight adult community mental health teams (CMHTs), and three CMHTs in Manchester.

The MINI deprivation score for Croydon varies widely (range 81.7 to 111.1) within what is overall an area with average levels (mean MINI: 100.1) of deprivation for England, making Croydon a nationally representative location. Specialist mental health services are organised into three localities containing eight CMHTs with each CMHT having a multidisciplinary mix of professionals, and access to beds. Croydon Borough is not inner-city, and not predominantly multi-ethnic. The three teams from Manchester were chosen to add these characteristics to the sample frame. The participating teams had MINI scores of 115 (two teams) and 125, the highest indicators for mental health needs outside London, with multi-ethnic populations and limited primary care mental health services.

### *Participants*

The inclusion criteria for general practices was that they provided care for residents in the catchment area of the participating teams.

### *Measures*

The Threshold Assessment Grid (TAG) is a one-page referrer-rated assessment of mental health problem severity<sup>10</sup>. It is completed by making one tick to indicate level of severity in each of 7 domains: (i) intentional self-harm; (ii) unintentional self-harm; (iii) risk from others; (iv) risk to others; (v) survival needs/disabilities; (vi) psychological needs/disabilities; and (vii) social needs/disabilities. The scale is “None”, “Mild”, “Moderate” and “Severe” (4-point scale, ranging from 0-3) for domains (ii), (iii), (vi) and (vii), with an extra “Very Severe” domain (score 4) possible for the remaining 3 domains (which may require

immediate action). The TAG total score is the sum of the seven items, and ranges from 0 (least severe mental health problems) to 24 (most severe). TAG was developed using innovative consensus techniques: six search workshops (n=57) followed by a Delphi Consultation (n=58)<sup>10</sup>. The psychometric properties were then investigated in referral cohorts to ten adult and older adult mental health teams (n=605), showing good construct and concurrent validity, internal consistency and test-retest reliability, and adequate inter-rater reliability<sup>13</sup>. It was also shown to be feasible for routine clinical use by primary care referrers to adult mental health services, with a cut-off TAG total score of 5 maximising sensitivity (76%) and specificity (50%) in matching mental health team view of suitability<sup>9</sup>. Further information is available from [www.iop.kcl.ac.uk/prism/tag](http://www.iop.kcl.ac.uk/prism/tag).

A customised version was developed for each study site, since tailored interventions are more likely to be implemented<sup>14</sup>. The TAG score sheet comprised just the one-page rating form, and together with a checklist of aspects to consider in rating the TAG<sup>10</sup> was used in Manchester. A TAG pack was used in Croydon which included the TAG score sheet, information about local mental health services, guidance on the referral letter contents<sup>15</sup> a checklist of aspects to consider in rating the TAG<sup>10</sup>, and a request to explicitly state for patients with a TAG total score of 4 or less for justification why they require referral to specialist mental health services.

A referral rating form for use by the CMHT was adapted from previous work<sup>9</sup>, and recorded: sociodemographic and clinical information; TAG scores (where applicable); the primary outcome of appropriateness of referral (appropriate versus not appropriate); and the secondary outcomes of ease of identifying urgency and ease of identifying the right professional group to undertake the initial assessment (both on a 5-point Likert scale, 1=Very

Difficult to 5=Very Easy) and discussion time. All outcomes are at the individual referral level.

### *Randomisation of practices*

Allocation on the basis of GP practice clusters was performed by an independent statistician (who was blind to the practice identity) using computer-generated random numbers on permuted blocks, stratified by site (Croydon versus Manchester) and practice list size tertiles (small: <3,500, medium: 3,501-6,000, large: >6,001).

### *Intervention*

The intervention was at the cluster level. All GPs working in **control group** practices were asked to continue with their usual referral practice: either a referral letter (Croydon) or use of an existing referral form (Manchester). All GPs working in **intervention group** practices were asked to continue with their existing method and in addition to complete and attach a TAG, using the TAG pack in Croydon or the TAG score sheet in Manchester.

### *Recruitment of Practices*

A Local Implementation Group (comprising key stakeholders from local services) provided advice on running the study. The Primary Care Trust, Local Medical Committee and local development meetings were involved / informed of the study as appropriate.

In the six months before the intervention began (in January 2005), letters were sent to all Practice Managers, with telephone follow-up by the researchers. An average of six telephone contacts per practice were required to either arrange a visit to the practice to discuss the study, to establish the practice was opting out. The study team provided lunch during the GP

practice visits to encourage attendance from as many GPs and other staff as possible. During the visit, the purpose of the study was explained, and the TAG pack (in Croydon) or TAG score sheet (in Manchester) were distributed. Questions were constructively replied to but no training was provided for completing the TAG, since it is intended to be used without formal training. Study awareness and retention were maximised by letters from the Mental Health Trust Borough Director before the study began (Croydon only) and distribution of a newsletter during the data collection phase.

### *Procedure*

The referral rating form was completed for all referrals received by the CMHTs between 1 January and 30 September 2005. It was completed when the referral was discussed, generally in the team meeting, *i.e.* solely on the basis of referral information, before CMHT contact with the patient. Local arrangements were made to ensure that it was also completed for emergency referrals processed by the CMHT outside of the team meeting. CMHTs were blind to allocation status, but blinding could not be fully maintained since referrals including a TAG were known to have come from an intervention group GP.

Costs data were collected to explore the cost resources of the intervention: (a) the number of GP contacts for referred patients was recorded for the six-months prior to referral and for the period since referral; and (b) anti-depressant and atypical antipsychotic prescribing costs incurred at primary and secondary care were collected for the 6 months before and after the study started. Data from the Primary Care Trust could be disaggregated by allocation status, whereas data from the secondary Mental Health Trust could not.

GPs from both intervention and control groups, mental health team leaders and consultant psychiatrists were invited to participate in semi-structured interviews to explore issues around the primary-secondary care interface, referral and use of TAG. Purposive sampling of GPs was used to ensure variation in practice size, and GP gender, ethnicity and experience. All interviews were audio-taped with consent and transcribed verbatim.

### *Analysis*

Stata version 8<sup>16</sup> was used. Outcomes were compared at follow up using chi-squared tests and independent sample t-tests. Secondary outcomes measured on 5-point scales were converted to binary variables for ease of interpretation and consistency with the primary outcome. An intention-to-treat (ITT) analysis was performed, followed by analysis of the appropriateness and ease of rating urgency and profession using a ‘complier average causal effect’ (CACE) analysis. This analysis is based on the assumption that the proportions of potential compliers are the same in both arms, and that the mean for non-compliers in the control arm is the same as that for the non-compliers in the intervention arm<sup>17</sup>. The effect being estimated is that of actually using the TAG rather than being given the opportunity to use it (as in an ITT analysis). ITT and CACE analyses were performed as follows: for binary variables, a generalised linear model which models probabilities directly; for continuous outcomes, linear regression; for rates, poisson regression and for the CACE analysis, instrumental variable analysis (two stage least squares). Stata commands *binreg*, *regress*, *poisson* and *ivreg* were used for these analyses respectively. In all cases site was included as a fixed effect and clustering by practice was included by using the *cluster* option, or in the case of *ivreg*, bootstrapping of clusters, a procedure which also corrects standard errors for misspecification of the distribution.

The cost of using the TAG was estimated using information on the material costs of producing and distributing the instrument and staff time spent completing and reading it. Primary care contact rates were compared using a poisson regression model with the time during which contacts could be made used as the exposure variable and the group variable as the independent variable. For follow-up GP contacts, the baseline rate was entered as an additional independent variable.

In the nested qualitative study, interviews with the primary care (GP) and secondary care (CMHT team leaders, consultant psychiatrists) participants were analysed thematically by constant comparison<sup>18</sup>. Themes emerging from one set of professionals informed the interview schedule used with the other. Analysis was completed independently by researchers with differing professional backgrounds (GP, psychiatry, nursing, psychology), with themes agreed through discussion<sup>19</sup>.

## **RESULTS**

### **Practices**

The trial flow diagram for allocation and involvement of GP practices is shown in Figure 1.

*Insert Figure 1 here*

Participating GP practices providing care for 408,839 patients (297,756 in Croydon, 111,083 in Manchester).

Four of the 27 non-participating practices gave no reason, and some gave more than one. Reasons for opting out were: 'Too busy' (n=12), 'Already have too much paperwork' (n=9),

‘Not interested’ (n=4), ‘Don’t see the benefit for the practice of the research’ (n=4), ‘TAG is too complicated’ (n=1) and ‘Not interested unless paid’ (n=1). Opt-out GP practices provided care for 119,961 patients (56,078 in Croydon, 68,883 in Manchester), comprising 6 large practices (4 Croydon, 2 Manchester), 6 medium practices (1 Croydon, 5 Manchester), and 15 small practices (5 Croydon, 10 Manchester). Hence small practices (<3,500 patients) comprised 54% of opt-out practices and 29% of participating practices. The opt-out rate was higher in Manchester, for reasons explored later in the results.

## **Patients**

1,061 referrals were made by 281 participating GPs to 11 CMHTs. The characteristics of the referred patients are shown in Table 1.

*Insert Table 1 here*

Researchers attended 451 (89%) of 508 referral meetings held during the study period. Of the 57 unattended meetings, data were provided by a team member for 52, indicating missing data from 5 (1%) referral meetings about 6 (1%) of patients. Outcome data were therefore available for 1,055 patients.

## **Outcomes of the intervention**

Implementation was low – the TAG was used with 25% of intervention group referrals (14% Manchester, 28% Croydon,  $P=0.07$  for difference between sites). There was no evidence for difference in the gender or diagnosis of patients with and without TAGs attached.

Table 2 shows the primary and three secondary outcomes by trial arm on an intention-to-treat basis for the 1,055 (99%) rated referrals. There were no significant differences at  $P=0.05$  between the two trial arms in any outcome. Logistic analyses controlling for site and practice or GP (included as random effects) showed no significant differences at  $P=0.05$  for any of the comparisons. The intraclass correlations for appropriateness among referrals from the same practice and also among referrals from the same GP were 0.06 in both cases.

CACE analysis indicated that the estimated mean primary outcome (appropriateness of referral) for compliers before using the TAG was 49%, for rating urgency was 60%, and identifying the professional was 87%. Table 2 includes the CACE analyses, showing greater (but still non-significant) treatment effects than the ITT analyses (for example, complier mean for the primary outcome increased by 16%).

*Insert Table 2 here*

### **Resource consequences**

Resource consequences were investigated in Croydon only. Thirty of the 55 Croydon GP practices supplies contact rates. Patients referred by intervention group GPs had higher post-referral primary care contacts rates than control group patients (Table 3). The difference after adjusting for differences in baseline contacts and time between referral and audit was statistically significant. These rates extrapolated over a one-year period result in an extra cost of £42pa for the intervention group (Curtis and Netten, 2005).

*Insert Table 3 here*



## **Implementation issues**

Reasons for the inadequate implementation were explored qualitatively with 35 GP referrers and 17 CMHT leaders and consultant psychiatrists. Two types of implementation block were identified: professional (for both referrer and referred-to team) and organisational. Illustrative data are shown in Figure 2.

*Insert Figure 2 here*

For GPs, forgetting to use the TAG when making a referral (as so few referrals were made that TAG use did not become routine) was not the only reason that TAG was not completed. GPs suggested that the TAG was simplistic and so did not reflect the complexity of dealing with patients with mental health problems. Some GPs expressed concern that the TAG score could be manipulated by other GPs to coerce the CMHT to accept referrals, and other GPs feared that TAG would be used by CMHTs to further restrict referrals.

For CMHT respondents, the view was expressed that GPs were neither willing to complete schedules nor reliable in their completion of TAGs. However, they also reported that TAGs accompanying referrals had not been considered in their referral meetings, so TAG scores had not in fact affected their decision-making.

At the organisational level, the two sites used differing approaches to implementation. In Croydon, the evaluation was called a service development, and directly supported by the mental health trust. In Manchester the evaluation was not part of a service change, so was perceived as research and practices were more able to initially refuse to participate in the

study and to later opt out of using TAG. This may account for a lower GP practice participation rate and lower use of TAG in Manchester.

## **DISCUSSION**

This multi-site multi-method cluster randomised controlled trial investigated the introduction of a standardised assessment of mental health problem severity into the referral process from primary to secondary care. The use of TAG did not impact on CMHT views about the ‘appropriateness’ of the referral, and so the intervention, whilst of a low cost, was not shown to be effective. The intervention was only implemented by 25% of intervention group GPs. The nested qualitative investigation in our study identified two barriers to implementation: professional (*e.g.* degree of trust, interpersonal relationships) and organisational (*e.g.* perception that standardised referral approaches are a camouflaged approach to rationing, differing perceptions about the importance of severity).

### **Strengths and limitations**

Strengths include sample size and design. The participating sample was large: 73 GP practices providing care for 408,839 patients, *i.e.* 0.8% of the population of England. The cluster design is an appropriate approach to minimise within-practice contamination. The use of a multi-method approach gives an understanding of why implementation was limited.

The study had several weaknesses. The impossibility of fully blinding participating CMHTs to allocation status points to the need for objective and unbiased measures of outcome, and the use of a subjective outcome of appropriateness as the primary outcome increases the possibility of bias. This outcome was chosen to have maximum validity, since the goal of the intervention was to improve agreement. The general issues of the low level of

implementation and the difficulty in using trial methodology to evaluate complex interventions are discussed below.

### **Managing the primary-secondary care interface**

Several barriers to accessing both primary and secondary mental health care have been identified, including no primary care contact, poor recognition, and in-system access barriers<sup>20</sup>. There is evidence of inequitable primary to secondary mental health care access, with non-clinical factors impacting on decisions about whether to refer, *e.g.* gender<sup>4</sup> and ethnicity<sup>21</sup>. When people are referred due to factors other than clinical need, an inefficient system results: unnecessary referrals are made, leading to reduced access for necessary referrals. The situation is further complicated by emerging evidence of the lack of benefit from early referral of common mental disorders over watchful waiting in primary care<sup>22</sup>. Active management of the primary–secondary care interface is needed to ensure equity of access.

Our study can inform efforts to improve primary–secondary care communication in two ways. First, caution should be exercised over the introduction of a new process such as a referral form. Prior to our study, the TAG had been carefully developed over a ten-year period within an externally funded research programme to develop a standardised mental health referral form. Four previous research grants had funded a systematic review, Delphi Consultations, expert consensus workshops, and a ten-site prospective cohort study evaluating the TAG. The rationale for its use was explained in our study through visits by researchers to 60 of the 72 participating practices. Since most new processes will be less tested before introduction and less explained when implemented, the likelihood of benefits arising may be even lower.

Second, the narratives of both GP referrers and referred-to team leaders and psychiatrists concentrated on the relationships between the health professionals, and how this influenced the referral process and outcome for both patient and professional. This indicates that formal referral processes (*i.e.* the paperwork) are embedded in a rich interpersonal context<sup>6,8</sup>. More attention to these processes (such as identifying and minimising implementation barriers) in the TAG development phase (rather than our focus on psychometric properties and feasibility of the measure itself) might have led to a higher rate of implementation. Future research will need to use methods which investigate formal process changes as only one part of a multi-level intervention to improve communication and mutual understanding across the interface.

### **Treatment fidelity in complex interventions**

Our study raises a general methodological issue. The uptake of the intervention was low, so uncertainty remains about its effectiveness<sup>23</sup>. In other words, even a rigorous trial based on current best practice in complex intervention evaluation<sup>24</sup> may not yield clear-cut results. Inclusion of a ‘process evaluation’ – collection of information to understand how the intervention is implemented and received<sup>25</sup> – indicated the relevance of multiple contextual factors.

The importance of context for complex interventions is becoming apparent<sup>26</sup>. Interventions which are tailored to the setting are more effective<sup>14</sup>, but varying the intervention conflicts with the goal of minimising variation in intervention implementation (*i.e.* treatment fidelity). This tension can be addressed using two approaches.

The first approach to considering context involves amending the intervention to the minimal degree necessary to allow implementation in each site. Our study used this approach, by having TAG either as a stand-alone single-page adjunct to an existing referral form, or a multi-page elaborated pack given to referrers. However, we showed that not just the setting but multiple contextual aspects from interpersonal relationships between individual participants to organisational beliefs were relevant to implementation. Our study design took no account of the impact of these moderators<sup>27</sup> for example using different approaches with GPs who were more or less favourable towards the importance of severity, who had or did not have an existing positive relationship with their CMHT, *etc.*

The second approach to context involves treating the intervention as a collection of options all based on a single coherent theoretical base. This means that the content of one implementation of the intervention may overlap totally, partially or minimally with another instance. This approach has been used with patient-level interventions, such as the development of manuals for psychological therapies. It has not been used with service-level interventions, perhaps because of the difficulty in describing their theoretical basis. In our study, this might involve the use of TAG as one of several elements of an overarching package of interventions to improve primary–secondary care communication. The higher proportion of small list size practices who opted out may point to the need for a different type of intervention, depending on practice size.

The complexity of complex interventions lie on a continuum. Those at the more complicated end are concerned with services or systems rather than patients, address problems characterised by polarised disagreement (*e.g.* how important is severity in deciding to refer?), and require attitudinal change for implementation. There is a need for methodological

development to combine the strengths of clinical trials with a recognition of this contextual complexity. One approach is to modify standard trial designs to account for patient preference, although systematic differences may exist between patients expressing or not expressing a preference, or between those willing or unwilling to be randomised<sup>28</sup>. Where an adequate theoretical basis can be established, further modifications to trial design may be needed to investigate systematically varied interventions. For complex interventions in complicated contexts, it may however become necessary to employ evaluative methodologies which treat context as an opportunity rather than a threat, such as the realistic evaluation approach to investigating how mechanisms acting in contexts produce outcomes<sup>29</sup>.

**Funding**

The study was funded by NHS Service Delivery and Organisation (SDO) Research and Development Programme (SDO/71/2003). The views expressed in this paper do not represent those of the funders.

**Ethical approval**

This study was approved by Metropolitan Multi-centre Ethics Committee (04/MRE11/8) with Local REC approval in London and Manchester, and research governance support from the Manchester and Croydon PCTs and South London and Croydon Maudsley NHS Trust and Manchester Mental Health and Social care Trust.

**Acknowledgements**

The study was funded by the NHS Service Development and Organisation R&D Programme. We are grateful to the participating primary care and mental health service staff in Croydon and Manchester.

**Declaration of interest**

None.

## References

1. Verhaak, P.F., Van De Lisdonk, E.H., Bor, J.H. & Hutschemaekers G.J.M. (2000). GPs' referral to mental health care during the past 25 years. *British Journal of General Practice* **50**, 307-308.
2. King, C. (2001). Severe mental illness: Managing the boundary of a CMHT. *Journal of Mental Health* **10**, 75-86.
3. Vázquez-Barquero, J.L., Herran, A., Simon, J.A. (1999). Epidemiology of mental disorders in the community and primary care. *Common mental disorders in primary care*. (Ed. Tansella, M. & Thornicroft, G). Routledge: London. 3-16.
4. Ross, S., Moffat, K., Mcconnachie, A., Gordon J. & Wilson, P. (1999). Sex and attitude: a randomized vignette study of the management of depression by general practitioners. *British Journal of General Practice* **49**, 17-21.
5. Evans, A., (1993). A study of the referral decision in general practice. *Family Practice* **10**, 104-110.
6. Chew-Graham CA, Slade M, Montana C, Stewart M, Gask L. (2007) A qualitative study of referral to community mental health teams in the UK: exploring the rhetoric and the reality. *BMC Health Services Research*, **7**, 117.
7. Harrison, J., Kisely S.R., Jones J.A., Blake, I. & Creed, F. H. (1997) Access to psychiatric care. *Journal of Public Health Medicine* **19**, 69-75.
8. Chew-Graham CA, Slade M, Montana C, Stewart M, Gask L. (2008) The loss of doctor-to-doctor communication: Lessons from the reconfiguration of mental health services. *Journal of Health Services Research & Policy*, **13**, 6-12.
9. Slade, M., Cahill, S., Kelsey, W., Leese, M. & Powell, R. (2002a). Threshold 4: an evaluation of the Threshold Assessment Grid as an aid to mental health referrals. *Primary Care Mental Health* **1**, 45-54.



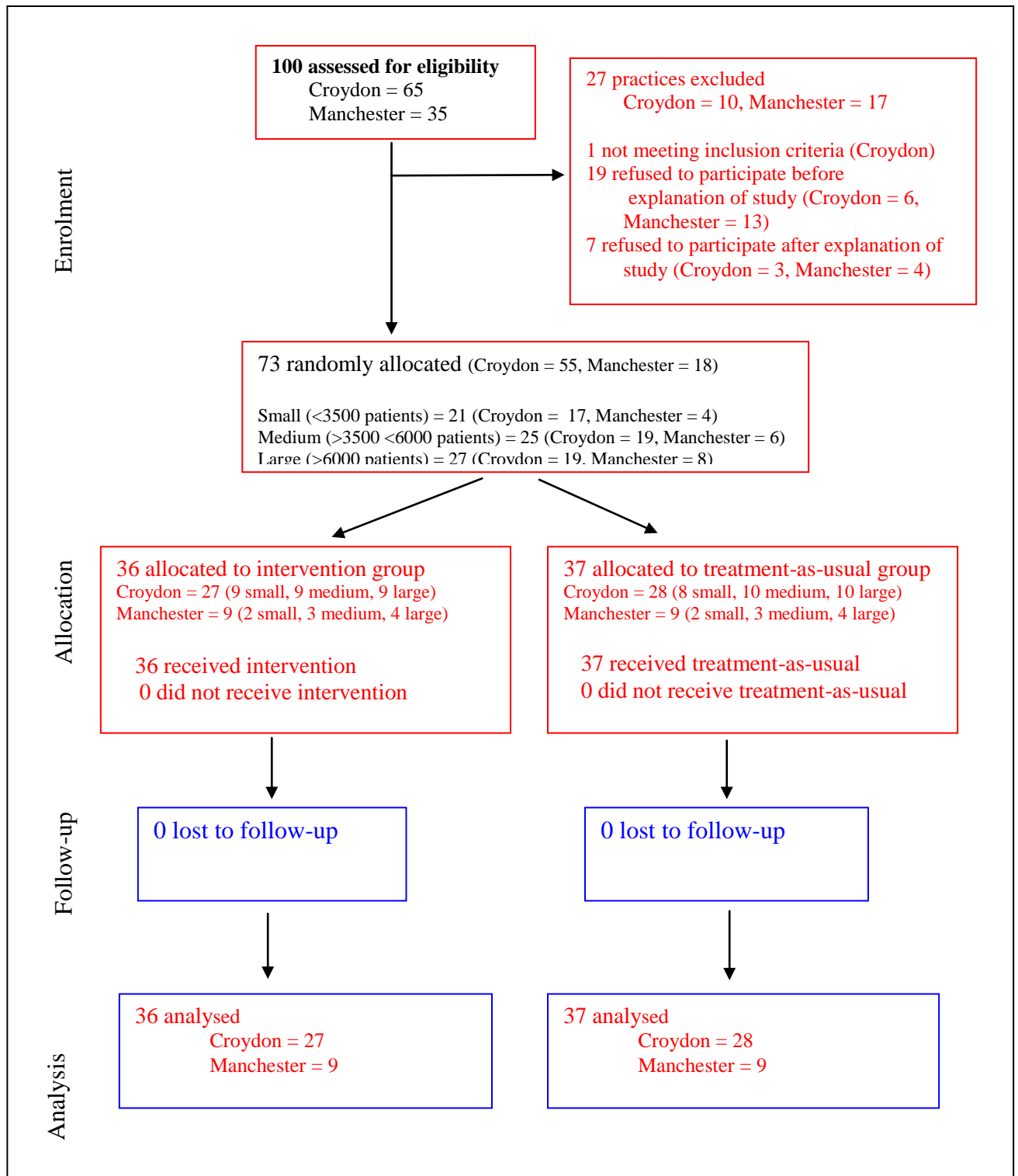
10. Slade, M., Powell, R., Rosen, A. & Strathdee, G. (2000) Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* **35**, 78-85.
11. Medical Research Council. (2002.) *Cluster randomised trials*. MRC: London.
12. Slade, M., Powell, R., Strathdee, G. (1997) Current approaches to identifying the severely mentally ill. *Social Psychiatry and Psychiatric Epidemiology*, **32**, 177-84.
13. Slade, M., Cahill, S. & Kelsey, W. (2002b). Threshold 2: The reliability, validity and sensitivity to change of the Threshold Assessment Grid (TAG). *Acta Psychiatrica Scandinavica* **106**, 453-460.
14. Lomas, J. (1993). Diffusion, dissemination and implementation: who should do what? *Annals of the New York Academy of Science* **703**, 226-257.
15. Strathdee, G,. (1990). Delivery of psychiatric care. *Journal of the Royal Society of Medicine* **83**, 222-225.
16. Statacorp. (2005). *Stata version 8.2*. Tx: College Station.
17. Dunn,. G, Maracy,. M, Dowrick, C., Ayuso-Mateos, J.L., Steffen Dalgard, O., Page, H., Lehtinen, M.D., Casey, P., Wilkinson, C. (2003). Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow up. *British Journal Psychiatry* **183**, 323-331.
18. Strauss, A. (1987). *Qualitative analysis for social scientists*. Cambridge: Cambridge University Press.
19. Henwood KL, Pidgeon NF. Qualitative research and psychological theorizing. *British Journal of Psychology* 1992, 83: 97-111.
20. Rosen, R., Florin, D. & Dixon, J. (2001). *Access to Health Care – Taking Forward the Findings from the Scoping Exercise*. NCCSDO: London

21. Shepherd, M., Gunnell, D., Maxwell, B. & Mumford, D. (1998) Development and evaluation of an inner city mental health team. *Social Psychiatry and Psychiatric Epidemiology* **33**, 129-135.
22. Kendrick, T., Simons, L., Mynors-Wallis, L., Gray, A., Lathlean, J., Pickering, R., Harris, S., Rivero-Arias, O., Gerard, K., Thompson, C. (2006) Cost-effectiveness of referral for generic care or problem-solving treatment from community mental health nurses, compared with usual general practitioner care for common mental disorders. *British Journal of Psychiatry*, **189**, 50-59.
23. Rychetnick, L., Frommer, M., Hawe, P. & Shiell, A. (2002). Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health* **56**, 119-127.
24. Campbell, J., Fitzpatrick, R., Haines, A., Kinmonth, A.L., Sandercock, P., Spiegelhalter, D. & Tyrer, P. (2000.) Framework for design and evaluation of complex interventions to improve health. *British Medical Journal* **321**, 694-696.
25. Oakley, A., Strange, V., Bonell, C., Allen, E. & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal* **332**, 413-416.
26. Campbell, N.C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., Guthrie, B., Lester, H., Wilson, P., Kinmonth, A.L. (2007) Designing and evaluating complex interventions to improve health care. *BMJ*, **334**, 455-459.
27. Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* **51**, 1173-1182.
28. King, M., Nazareth, I., Lampe F., Bower, P., Chandler, M., Morou, M. Sibbald, B. & Lai, R. (2005) Conceptual framework and systematic review of the effects of participants'

and professionals' preferences in randomised controlled trials. *Health Technology Assessment* **9** (35), 1-186.

29. Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. Sage: London.

**Figure 1: Trial flow diagram for GP practice involvement**



**Figure 2: Qualitative exploration of TAG implementation blocks from the perspectives of GPs and Community Mental Health Team leaders and psychiatrists**

Theme	Illustrative data
<b>Professional factor – referrer perspective</b>	
Simply forgetting	<ul style="list-style-type: none"> <li>• <i>I don't know where these TAG forms are in my practice. [GP 192]</i></li> <li>• <i>I have to admit I've not remembered it every time...but also I can go for months and months without making a referral. [GP 649/176]</i></li> </ul>
Perception of TAG as simplistic/reductive	<ul style="list-style-type: none"> <li>• <i>...it does ask the question about why you need to have a sort of score sheet in the first place. If I can't deal with it, I can't deal with it... [GP 660/175]</i></li> </ul>
Fear of how TAG was used by the CMHT	<ul style="list-style-type: none"> <li>• <i>...But the way to do it is not to make the hoops ever more difficult to jump through because GPs are world class, if there was a hoop jumping Olympics we'd flippin' win hands down, yeah? Making it more difficult to get a referral through, it will not stop inappropriate referrals. [GP616/191]</i></li> </ul>
Suggestion that TAG could be manipulated	<ul style="list-style-type: none"> <li>• <i>...You get good at ticking boxes, yeah, I'm not going to set myself up to fail for my patients, I'll advocate for my patients so I'll tick the boxes. "Oh, oh, they said they were going to kill themselves to me". [GP 616/191]</i></li> </ul>
<b>Professional factor – CMHT perspective</b>	
TAG not used in referral meetings	<ul style="list-style-type: none"> <li>• <i>I don't get a sense that it, it's affected our decision in any way, shape or form. [CMHT 10/7]</i></li> <li>• <i>No, we didn't [use TAG] no, I don't think we even looked at it, are we supposed to? [CMHT]</i></li> </ul>
Perception that TAG does not make a difference	<ul style="list-style-type: none"> <li>• <i>It doesn't actually change the outcome from the GP's point of view. [CMHT 23/1.12]</i></li> </ul>
Perception that GPs don't use/can't use	<ul style="list-style-type: none"> <li>• <i>....GPs using rating scales is perhaps unfamiliar territory, I don't know. ...they might not have felt confident in doing that more formal assessment and having to put something to paper which they could later be taken up on. [Psych 1]</i></li> </ul>
Perception that GPs give different information on TAG than on referral letter	<ul style="list-style-type: none"> <li>• <i>...and what sometimes is quite interesting is that the letters that are sent doesn't quite correlate to what the TAG says... [CMHT 17.7]</i></li> </ul>
<b>Organisational factors</b>	
Mental health problems are complex	<ul style="list-style-type: none"> <li>• <i>...and I think that people are not as clear cut as numbers. So there may be people who could really benefit from seeing, or having, a CPN, but maybe are not scoring that highly. And if it's a case of using the resources effectively...I know we have to do that. [GP 357]</i></li> <li>• <i>Somehow it doesn't always translate though where, the boxes to patient. You don't really get a, a true feel for the client somehow. [PSYCH 23/3/7]</i></li> </ul>
Interface issues	<ul style="list-style-type: none"> <li>• <i>...yet another form to fill when we really need to overhaul the whole system [GP 660/175]</i></li> <li>• <i>...I think the Tag could be very easily manipulated to up the ante. We do it all the time to try and get a</i></li> </ul>

---

Perception of role of CMHT	<p><i>bed. You know, well we would do. If you thought, you'd try and get that score as high as you can. [CMHT 24/1.14]</i></p> <ul style="list-style-type: none"> <li>• <i>.... but there are situations where we see a patient on a bad day, then they're assessed by the team on a slightly better day, but it's a fluctuating scenario. And its almost as if you get a snapshot. Oh and you know I suppose the cynical view is "oh that's a relief. This patient isn't really appropriate for us, so we don't need to be involved". Whereas in fact the next day the person could be really chaotic again....And I think its partly a feature of the service as a whole, is that those people are not well served. [GP630/168]</i></li> <li>• <i>...I don't know how much they take the scoring into account if its still not a severe and enduring mental health problem, I don't know how they would relate that together... [GP 643/165]</i></li> <li>• <i>...how do you define what severe mental illness? You know, somebody may have schizophrenia but be really well maintained. Somebody may have, you know, mild depression and anxiety and be creating havoc. [PSYCH 23/3/8]</i></li> </ul>
----------------------------	--

---

**Table 1: Referral information for patients referred by GPs to CMHTs (n=1,061)**

		<b>Total</b>			<b>Croydon</b>		<b>Manchester</b>	
		<i>Overall</i> n=1,061	<i>Intervention</i> n=517	<i>Control</i> n=544	<i>Intervention</i> n= 379	<i>Control</i> n= 455	<i>Intervention</i> n=138	<i>Control</i> n=89
<b>Gender</b> <sup>1</sup>	Female n (%)	578 (55)	280 (55)	298 (55)	208 (56)	250 (55)	72 (53.6)	48 (54)
<b>Age</b> <sup>2</sup>	Mean (s.d.)	36.2 (12.1)	36.2 (12.1)	36.3 (12.1)	36.5 (12.2)	36.5 (12.2)	35.3 (11.8)	35.3 (11.8)
<b>Clinical Diagnosis</b>	n(%)							
	Psychosis/Schizophrenia	93 (9%)	48 (9%)	45 (8%)	30 (8%)	34 (8%)	18 (13%)	11 (12%)
	Anxiety Disorder	173 (16%)	81 (16%)	92 (17%)	73 (19%)	88 (19%)	8 (6%)	4 (5%)
	Depressive Disorder	478 (45%)	243 (47%)	235 (43%)	197 (52%)	210 (46%)	46 (33%)	25 (28%)
	Bipolar Disorder	32 (3%)	18 (3%)	14 (3%)	14 (4%)	12 (3%)	4 (3%)	2 (2%)
	Other	146 (14%)	54 (11%)	92 (17%)	48 (13%)	87 (19%)	6 (4%)	5 (6%)
	Unknown/Missing	139 (13%)	73 (14%)	66 (12%)	17 (4%)	24 (5%)	56 (41%)	42 (47%)

<sup>1</sup> 5 missing ratings

<sup>2</sup> 3 missing ratings

**Table 2: Comparison of mental health team ratings made about referrals from intervention and control group practices (n=1,055)**

	Intervention N=514	Control N=541	ITT effect, adjusted for site and clustering by practice		Complier average effect adjusted for site and clustering by practice	
			Difference (Intervention- Control) (95% CI)	P	Difference (Intervention- Control among compliers) (95% CI)	P
<b>Appropriateness of referral for the team</b> <i>n (%) of ratings Appropriate</i>	330 (64%)	326 (60%)	4% (-5% to 11%)	0.41	13% (-22% to 49%)	0.46
<b>Ease of identifying urgency of referral</b> <i>n (%) of ratings Easy or Very Easy</i>	277 (81%)	253 (76%)	4% (-2% to 10%)	0.15	18% (-10% to 45%)	0.22
<b>Ease of identifying the appropriate profession for the initial assessment</b> <i>n (%) of ratings Easy or Very Easy</i>	303 (89%)	292 (87%)	2% (-3% to 8%)	0.46	6% (-21% to 33%)	0.65
<b>Time (minutes) to discuss referral (n=646)</b> <i>mean (s.d.)</i>	3.65 (2.08)	3.81 (2.15)	-0.15 (-0.48 to 0.18)	0.37	-0.52 (-1.93 to 0.89)	0.47



**Table 3: Service use data in Croydon**

Type of data	Intervention	Control	Incidence rate ratio <sup>1</sup> (95% CI)	P
<b>Primary care contacts with referred patients</b>				
Number of practices supplying data	16	14		
Number of referred patients	206	178		
Primary care contact rate (s.d.) in 6 months before referral	4.7 (3.6)	4.3 (3.6)	1.09 (0.90 to 1.32)	0.375
Equivalent annualised rate	9.4 (7.1)	8.6 (7.1)		
Primary care contact rate (s.d.) between referral and audit	4.8 (5.3)	3.4 (3.1)	1.36 (1.08 to 1.70)	0.008
Equivalent annualised rate	11.3 (10.2)	8.8 (7.9)		
<b>Prescription costs (£)</b>				
<i>Primary Care Trust SSRI costs</i>				
Six months before trial	155,461	195,637		
Six months after trial	139,856	176,945		
<i>Primary Care Trust atypical antipsychotic costs</i>				
Six months before trial	155,739	247,708		
Six months after trial	142,976	234,089		
<b>Combined</b>				
<i>Mental Health Trust SSRI costs</i>				
Six months before trial <sup>2</sup>	28,766			
Six months after trial	22,753			
<i>Mental Health Trust atypical antipsychotic costs</i>				
Six months before trial	426,872			
Six months after trial	394,153			

<sup>1</sup>Adjusted for site

<sup>2</sup>Data missing for July and August 2004 so 4-month total multiplied by 1.5